On empirical generalisations

Federica Russo

Philosophy, University of Kent

f.russo@kent.ac.uk

Draft of March 7, 2011

Abstract

Manipulationism holds that information about the results of interventions is of utmost importance for scientific practices such as causal assessment or explanation. Specifically, manipulation provides information about the stability, or invariance, of the (causal) relationship between (variables) X and Y: were we to wiggle the cause X, the effect Y would accordingly wiggle and, additionally, the relation between the two will not be disrupted. This sort of relationship between variables are called 'invariant empirical generalisations'. The paper focuses on questions about causal assessment and analyses the status of manipulation. It is argued that manipulationism is trapped in a dilemma. If manipulationism is read as providing a conceptual analysis of causation, then it fails to provide a story about the methods for causal assessment. If, instead, manipulationism is read as providing a method for causal assessment, then it is at an impasse concerning causal assessment in areas where manipulations are not performed. Empirical generalisations are then reassessed, in such a way that manipulation is not taken as methodologically fundamental. The paper concludes that manipulation is the appropriate tool for some scientific (experimental) contexts, but not for all.

1 Introduction

Manipulationist theorists, in slightly different ways, hold the view that information about the results of interventions is of utmost importance for scientific practices such as causal assessment or explanation.[1]

More specifically, the information that manipulation is meant to provide concerns the stability, or invariance, of the (causal) relationship between (variables) X and Y. This means that, put in non-technical terms, were we to wiggle the putative cause X, the putative effect Y would accordingly wiggle and, additionally, the relation between the two will not be disrupted. This does not entail that wiggling X will necessarily make Y wiggle, but that, if it does, we will be interested in whether the relationship between X and Y is invariant in the sense sketched above. Relationships of this sort are called *invariant empirical*

---

[1]The main theoriser and partisan of the manipulatinist (or interventionist) account is no doubt Jim Woodward (see Woodward (2002, 2003, 2010)). The approach has been also endorsed and used for various purposes by many other scholars, for instance Baumgartner (2009); Glennan (2010); Hausman (1997); Hausman and Woodward (2004); Waters (2007); Woodward and Hitchcock (2003).

generalisations and have the characteristic of being exploitable for explanation or for causal assessment. In this paper, I focus on questions related to causal assessment rather than explanation. That is to say, I will be concerned with what makes empirical generalisations *causal* rather than with what makes them explanatory.

In §2, I present the manipulationist account of empirical generalisations. I make it clear that *manipulation* is central for the account. The rest of the paper investigates the *status* of manipulation for questions of causal assessment.

In §3 I argue that the manipulationist account is trapped in a dilemma. If the project is read as contributing to the conceptual analysis of causality, then it is at an impasse concerning the *methods* for causal assessment, i.e. no story about *how* to establish whether X causes Y is offered. If the project is read as contributing to the methodology of causality, then a second dilemma opens up. Strictly interpreted, manipulationism fails to offer methods for causal assessment in scientific areas where manipulations are not performed. Charitably interpreted, instead, manipulationim becomes so vague as to be an unilluminating—and even misleading—rationale underpinning causal reasoning in both experimental and nonexperimental contexts.

In the light of the previous discussion, in §4 I reassess empirical generalisations. The core of agreement with manipulationist theorists is that empirical generalisations are indeed change-relating relations and that for empirical generalisations to be causal they indeed have to be invariant, albeit in a sense that does not take manipulations as methodologically fundamental. The importance of the *change*-relating character of empirical generalisation has to do with the rationale underpinning causal reasoning: it is not manipulation but *variation* that does this job.

2 Manipulationist Empirical Generalisations

To understand the manipulationist project, we need to spell out the notions of (i) empirical generalisation, (ii) invariance, (iii) intervention, and the relations they stand with respect to each other.

An *empirical generalisation* is a relation between variables that has the characteristic of being change-relating or variation-relating: changes in the putative causal-variable X are associated with changes in the putative effect-variable Y. Of course, the problem of distinguishing causal from spurious or accidental generalisations immediately arises. We could hit upon a change-relating relation that is accidental: an increased number of storks might be statistically associated with an increased number of births, but arguably there is no causal link between these two variables. Or, a change-relating relation might be spurious: yellow fingers might be statistically associated with lung cancer but this is the case because they are effects of a common cause, that is cigarette smoking.

Change-relating relations have to show a certain invariability or stability in order to be causal (or to be explanatory—this falls outside the scope of the paper). Briefly put, this means that empirical generalisations have to be invariant under intervention or under a sufficiently large class of changes.

There are two different types of changes we might be interested in. The first sort of changes involves background conditions, whereas the second sort of changes is exactly those that establish invariance. The former, also called

'resilience' by Skyrms (1980) or 'context-unanimity' by Eells (1991), establishes some kind of stability of a relationship across different background conditions. The latter, instead, requires manipulating the variables figuring in the relationship itself, and we will call the generalisation invariant, roughly speaking, if changing values of the cause-variable changes values of the effect-variable, and yet, the relationship between the cause and effect-variables is not disrupted. This is the sort of *invariance* required for empirical generalisations to be causal. Invariant generalisations are then used to ask counterfactual questions about what would happen to the effect, had the cause been different.

However, in order to evaluate the effects of manipulations, not all counterfactuals will do. Relevant counterfactuals are those that describe outcomes of *interventions*. Consider an empirical generalisation between X and Y. An intervention I on X has to have three characteristics: (i) the change in the value of X is totally due to the intervention; (ii) the intervention will affect the value of Y, if at all, just through the change in the value of X; (iii) the intervention is not correlated with other possible causes of Y (other than X). Interventions serve to establish whether changes in the cause will bring about changes in the effect, and yet the relation between the cause and the effect remains unaltered. If this is the case, then invariant empirical generalisations are in fact causal in character.

Here is a stock example from physics (discussed, inter alia, in Woodward (2003)). Consider the ideal gas law, which states that the state of an amount of gas is determined by its pressure P, volume V, and temperature T: $PV = nRT$, where n is the number of moles of gas and R the universal gas constant. This empirical generalisation is invariant under a whole range of interventions on the temperature. Therefore the generalisation correctly describes how manipulations or interventions on the temperature of gas would affect the gas pressure, holding fixed the gas volume. Because this empirical generalisation is invariant, it is *potentially* exploitable for manipulation and control. In fact, we do not need to *actually* perform such interventions, but if we did, then we could see how the effect-variable would change upon interventions on the cause-variable, and the relation between the cause- and effect-variable remain stable.

Here is an illustration from biology. Consider Dawkin's fictious gene R (discussed in Woodward (2010)): when variant r is present, individuals have dyslexia and are unable to learn to read; when variant r' is present, individuals can learn and read normally. The relation between the gene R and the ability to learn and read is not a very stable one, however. In fact, differences in background conditions (such as schooling or culture), will disrupt the relation between R and learning and reading. In other words, outcomes in learning and reading are not dependent on manipulations on the gene R. This has to be contrasted, instead, with more stable relationships involving other genes, for instance for eye colors or for external sexual characteristics.

In sum, empirical generalisations can be exploited for explanation or causal assessment insofar as they are invariant under intervention. Causal assessment, in particular, amounts to asking counterfactual questions about the outcomes of interventions of the following sort: were we to manipulate the putative cause X, the putative effect Y would accordingly change, and the relation between the two would not be disrupted as a consequence of the manipulation on X. In the remainder of the paper, I investigate the *status* of manipulation for causal

assessment.

3   The Dilemma

The manipulationist account, I now argue, is caught in a dilemma. The dilemma arises because the manipulationist account can be given two readings: conceptual and methodological. *First*, according to the conceptual reading, the account aims to provide identity conditions for causal relations; the sought solution is that X causes Y if, and only if, manipulations on X would accordingly yield changes to Y. *Second*, according to the methodological reading, were manipulations on X to yield changes on Y, then we would be entitled to infer that X causes Y.[2]

Those two readings of manipulationisms lead to a dilemma. In the *first* case—that is if the project is conceptually read—manipulationism turns out to be unilluminating as to the methods to use for causal assessment. In the *second* case—that is if the project is methdodologically read—then a second dilemma opens up: (a) strictly interpreted, methodological manipulationism is not in a position to offer a solution in domains where it is not possible to intervene (typically, the social sciences, but also astronomy); (b) charitably interpreted, the requirement of manipulation becomes so vague as to be not only unilluminating, but also misleading, as to the rationale underpinning causal reasoning.

Ultimately, the dilemma mirrors a more profound problem in the philosophy of causality: the relation between epistemology/methodology and metaphysics. Two remarks are in oder. First, specific questions (epistemological/methodological and metaphysical) ought not to be conflated, and instead call for appropriate and distinct answers. That is to say, we should not give a metaphysical answer to a methodological question, and vice-versa. Second, it is vital to investigate how metaphysical issues have a bearing on epistemological and methodological ones, and vice-versa. However, this can be done only insofar as different types of questions and of answers are kept distinct. With these caveats in mind, let us now analyse the horns of the dilemma.

3.1   Horn 1: Conceptual Manipulationism

According to the *conceptual* reading of manipulationism, X causes Y *if*, and only if, manipulations on X accordingly yield changes to Y. This, notice, amounts to giving identity conditions for causal claims, and consequently the project contributes to the analysis of the concept of causation. Manipulation is here the *concept* in terms of which causation is cashed out. Under this reading, manipulationism says what has to be true *if* X causes Y.

If this reading is correct, then manipulationism is unilluminating as regards the *methods* to establish whether in fact X causes Y. Nevertheless, it is desirable to have a conceptual analysis of causation that goes hand in hand with methodology. Once we know what a causal relation between X and Y amounts

---

[2]Another strong proponent of the quandary above is Strevens (2007, 2008), who distinguishes between conceptual manipulationism and explanatory manipulationism. Given that I focus on questions of causal assessment rather than explanation, I prefer using the more general term 'methodological manipulationism'.

to, it helps a great deal to know *how* to find out what causes what. Conversely, if conceptual analysis and methodology are entirely disconnected, then our understanding and practice of causal inference are too fragmented to be successful. Many objections to standard accounts of causation (probabilistic, counterfactual, regularity, interventionist) stem from the fact that (i) epistemological, methodological and metaphysical questions are conflated and that (ii) most often the bearing of the epistemology/methodology side on the metaphysics side (and vice-versa) have not been thoroughly investigated (on this point, see also Cartwright (2007)).

Thus, once we endorse the idea of having coherent (rather than disconnected) methodological and conceptual accounts of causation, then the only possible methodological candidate, under the conceptual reading of manipulationism, is a methodology based on manipulations. In this case, we have to investigate Horn 2 below, which discusses precisely methodological manipulationism.

It is worth noting that manipulationist theorists (and particulary Woodward) claim that the project is methodological rather than conceptual. Thus, Horn 2 below appears to be most relevant, at least prima facie. Nervertheless, as the discussion of Horn 2 will reveal, the objections to the methodological reading do press the manipulationist theorist back into Horn 1, whence its relevance for our purposes.

Yet, if escaping Horn 1 leads to Horn 2, and in turn, the branches of Horn 2 loop back into Horn 1, then it seems that the manipulationist is stuck between a rock and a hard place. But there is a way out: my reassessment of empirical generalisations offered in §4.

3.2   Horn 2: Methodological Manipulationism

According to the methodological reading, the perspective is rather reversed: were manipulations on X to yield changes on Y, *then* we would be entitled to infer that X causes Y. Manipulation is here a *method* to establish whether X causes Y. There is another dilemma opening up now. The requirement of manipulation can be either (a) *strictly* interpreted, or (b) *charitably* interpreted.

Horn (a): The strict interpretation

Strictly interpreted, manipulationism prescribes the following. In order to know whether X causes Y, perform an intervention on X, hold fixed anything else, and see what happens to Y.

The typical situation where this happens is the controlled experiment. Put it in simple terms, in a controlled experiment we compare results obtained from two groups: the experimental and the control group. Those are similar in all relevant respects, except for the putative cause, which undergoes manipulation in the experimental but not in the control group. The experimenter can then assess the influence of the putative cause X on the putative effect Y (i) by holding fixed any other possible influence of factors other than X, and (ii) by varying *only* the putative cause. Thus, in a controlled experiment, manipulation is indeed the *key* tool to establish causal relations.

Let me make clear what the *status* of manipulation is. Manipulation is a *tool* to get to know what causes what. It is also worth noting that the controlled experiment is here oversimplified and all the difficulties of experimental

design are overlooked. Identifying the right or best groups (of people or any other type of units) to include in the experiment, choosing the right or best intervention to perform, and assessing the effects of such intervention are all far from being trivial and obvious tasks. Randomisation, *the* controlled experiment par excellence, is indeed simple *in principle*, but not *in practice*. In practice, experimental design is a complex and delicate thing, having its own research tradition tracing back at least to the seminal work of Fisher (1935).

The problem, however, is that most situations in social science (and some exist in natural science too, e.g. astronomy) are *not* like controlled experiments. In observational contexts in social science we need methods to find out what causes what *without* resorting to manipulation. The problem isn't new at all. Early methodologists in social science recognised this difficulty already in the Sixties. For instance, Blalock (1961), trained in mathematics and physics, promptly admitted that well-designed experiments *could* allow the scientist to make causal inferences based on the outcomes of manipulations "with some degree of confidence" and "with a relatively small number of simplifying assumptions" (ibidem, p. 4). Blalock then noticed that this isn't the case when scientists have to deal just with *observational* data. The question is not whether *in principle* the same rules of inference can be applied, but how practical difficulties can be overcome in order to make reliable causal inferences on the basis of data coming from nonexperimental studies.

This is not to say that the social sciences do not perform interventions at all. Policy interventions, for instance, are indeed interventions, but the status of manipulation is here different than the one in controlled experiments discussed earlier. Policy interventions are based on a causal story, i.e. on valid empirical generalisations. The results of policy interventions may then lead us to further confirm or to question the validity of the empirical generalisation. Thus, manipulation is *not* a tool to find what causes what. Instead, we 'manipulate' *because* we know what causes what (to the best of our knowledge). In other words, manipulation is a *consequence* of a causal story established (usually) in absence of interventions strictu sensu. Witness for instance Birkland (2010, p. 241):

> If the participants in policy making can at least approximate goal consensus, then the next thing they must do is to understand the *causal theory* that underlies the policy to be implemented. A *causal theory* is a theory about what causes the problem and what intervention (i.e. what policy response to the problem) would alleviate that problem. Without a good causal theory it is unlikely that a policy design will be able to deliver the desired outcome. [Emphases in the original.]

The manipulationist will then rebut that, to find out what causes what, we don't have to *actually* intervene—an ideal manipulation will do. In fact, the manipulationist thesis says that *were* we to intervene on the cause, the effect would accordingly change. Here are my replies to the objection.

First, some ideal interventions may not make any (physical) sense. For instance, imagining an intervention that would double the orbit of the moon (assuming Newtonian gravitational theory and mechanics) to see what would happen to the tides goes far beyond an ideal—in the sense of physically *possi-*

ble—intervention (Woodward, 2003, p.131). An intervention that is not physically possible—albeit ideal—must be *conceptual*. If we imagine moving the moon in a way that such and such changes on the tides will result, this spurs from our (already established) causal knowledge, but is not *evidence* to establish a causal relation between the moon and the tides.

Consequently, it is reasonable to suspect that the manipulationist project (also) has a conceptual flavour. This suspicion is reinforced by claims Woodward makes, such as "my aim is to give an account of the content or meaning of various locutions such as X causes Y" (Woodward, 2003, p.38). However, in this way, the manipulationist is in a loop that sticks her back into Horn 1 discussed earlier.

Second, some other ideal interventions cannot be tested and, therefore, causally evaluated. For instance, Morgan and Winship (2007, p.280) supported the argument that attributes such as gender can be *ideally* manipulated thus:

> [. . .] the counterfactual model could be used to motivate an attempt to estimate the average gain an employed black male working full time, full year would expect to capture if all prospective employers believed him to be white.

However there is no way of testing such 'thought-experiments' against real data. This, again, raises the suspicion that manipulationism is (also) a thesis about the *meaning* of causation. This, again, brings the manipulationist back to Horn 1.

Third, and most importantly, if the manipulationist stresses the counterfactual aspect of the thesis ('*were* we to intervene . . . '), then she is definitively providing a conceptual analysis of causation. The manipulationist is in fact stating what must be true about the relationship between X and Y, *if* X causes Y. 'What must be true' corresponds to providing the meaning—and therby a conceptual analysis—of 'cause' in locutions such as 'X causes Y'. Choosing whether manipulation, or other notions, supplies the meaning of 'cause' may depend on our a priori intuitions about causation as well as on an analys of the scientific practice—that's beyond the point at stake. The manipulationist theorist is indeed free to hold such a conceptual account, appealing to the best arguments she can produce. But the fact is, even if the conceptual story ('what must be true') is accepted, no methodological story ('how to know whether it is in fact true') is offered. Thus, the manipulationist is irriversibly brought back to Horn 1.

Horn (b): The charitable interpretation

Charitably interpreted, manipulationism does not prescribe that the *agent* intervenes to find out what causes what. If the agent cannot manipulate, *Nature* will will do it for us. Thus, once Nature has manipulated for us, causal assessment is about evaluating changes (or variations) in the putative effect Y due to changes (or variations) in the putative cause X. This is roughly what happens, for instance, in 'natural experiments' in economics or epidemiology. In these observational contexts the assignment of treatment is done 'by Nature' rather than 'by the experimenter'. Two remarks are immediately in order.

First, even if Nature can in principle manipulate (or randomise) for us, we need tools to find out *whether* Nature did in fact manipulate and, if so,

whether the manipulation was effective. This means, eventually, that we have to establish what causes what in *nonexperimental* situations. Notice that this is exactly the kind of impasse that resulted from the stric reading of Horn (a) discussed earlier. Invariance under *intervention* (strictly interpreted) then turns out to be too strong a requirement for causal assessment of empirical generalisations in nonexperimental contexts. No wonder, then, that the need for a weaker version of invariance, i.e. *not* based on manipulation, come from the quarters of manipulationists themselves: this is Woodward's notions of weak invariance and of possible-cause generalisations to be discussed later in §4.

Second, the charitable reading of methodological manipulationism suggests that what is of utmost importance is to evaluate whether *changes* in the putative effect Y occur as a consequence of *changes* in the putative cause X. If this is correct, then *manipulation* cannot be interpreted as providing the rationale underpinning causal reasoning. Such interpretation is misleading and disingenous. Let me elaborate this idea further.

If we let manipulation underpin causal reasoning, the risk is to create another 'gold standard', analogous to randomised clinical trials (RCTs) in evidence-based medicine. But RCTs aren't, by all means, the gold standard either. Criticisms of the alleged superiority of RCTs abound. Here is a notable one. Thompson (2011) is concerned that statistical methods *alone* cannot be a reliable tool for causal inference. More specifically, in his argument, the differences between trials in biomedical contexts and trials in agricultural settings—the origin of the Fisherian theory—are the key to understand why randomisation is by and large successful in the latter but not in the former.

The alleged superiority of manipulationist methods over observational ones is based on the idea that non-experimental models try to (actually, struggle to) reproduce the same methodology. Whence the widespread belief that experimental methods are intrinsically better than nonexperimental ones. This idea is however questionable. Each scientific method—be it experimental or observational—has its own virtues (and weakenesses). Consequently, the goodness of a method has to be evaluated in the context in which it is used. If we cannot manipulate, it makes no sense to say that a controlled experiment would have been better than a cohort study. What does make sense is, for instance, questioning whether the chosen sampling method for the cohort study at stake was good or not in the given context. Methods are to be evaluated for the job they are supposed to do, not with respect to an alleged gold standard.

The motivation to have a rationale underpinning causal reasoning is to unify different methods under a principle that embraces them all. However, manipulation cannot do that—the impasse that ensued from the strict reading (Horn (a)) made the point. But there is indeed *one* rationale that unifies manipulationist and observational methods as *methods for causal inference*: this is the rationale of *variation* that I shall present next in §4.

To sum up. Manipulationism can be read as contributing to a *conceptual* analysis of causation or as contributing to causal *methodology*. I argued that manipulationism is trapped in a dilemma because if the 'conceptual' reading is preferred, then the account is at an impasse with the methodology side. Conversely, if the 'methodological' reading is preferred, then either manipulationism is too strict and leaves out that part of science that finds out what causes what without manipulating, or it is too vague to be an illuminating rationale underpinning

causal reasoning.

4 Empirical Generalisations Reassessed

This section reassesses empirical generalisations. The arguments hereby presented are built upon the same formalism typically employed by manipulationist modellers: causal modelling (or structural modelling). *First*, I present causal modelling as the answer to the same methodological challenge identified by methodological manipulationism: how to find out what causes what. *Second*, I present the variational epistemology underpinning experimental and observational methods and I show how it works within causal modelling. *Third*, I develop a notion of invariance that does not necessarily require manipulation.

Let me make clear what the core of agreement with manipulationist theorists is. I do indeed share with them the idea that empirical generalisations are change-relating relations between variables of interest. I do also share the idea that for empirical generalisations to be causal, they have to be invariant, albeit in a sense that I will specify later and that does *not* necessarily involve the notion of manipulation. Let me anticipate the importance of characterising empirical generalisations as *change-relating*: this aspect reflects the *variational* epistemology that underpins causal modelling. The full argument is given below.

Causal modelling Causal modelling (also, or alternatively, called structural modelling) constitutes the common ground for discussion with manipulationist theorists.[3] Causal modelling is a methodology the purpose of which is to establish what causes what, in a given context. Causal modellers do so by specifying the data-generating process (or mechanism) that accounts for the observations in the data set. There is no need to go into technical details, especially related to the statistical properties and tests of those models. The interested reader is directed to Mouchart and Russo (2010); Mouchart et al. (2009); Russo (2009, 2011a); Russo et al. (2008); Wunsch (1988, 2007); Wunsch et al. (2010), besides the well-known works of e.g. Pearl (2000) and Woodward (2003).

Causal modelling can be schematically presented as step-wise methodology. The *first* step is to define the research question, the population of reference, and the causal context, broadly conceived. This includes taking into account well-established theories, comparative analyses, and preliminary analyses of data. Here is an example from social science in practice. Gaumé and Wunsch (2010) investigate the determinants of self-rated health (i.e. of the individual's subjective perception of his/her own overall health). The first thing they do is to specify their research question and to define, consequently, the population of reference and the context: they analyse data related to Baltic countries in the Nineties, that is in a post-communist socio-political context.

---

[3]In the literature, causal and structural modelling are used intercheangeably. I am not opposed to this practice, albeit a distinction between the two exists. 'Causal modelling' was introduced by methodologists such as Blalock in the Sixties, and covered different quantitative methods in social science. 'Structural (equation) modelling' is instead a term more familiar to econometricians, who intended to represent, with structural equations, the 'structure' of phenomena as prescribed by economic theory. The two terms can be legitimately used as synonyms insofar as causal models model mechanisms, that is causal structures (Russo, 2011a,b).

On the basis of the outputs of step one, the *second* step is to give structure to the joint probability distribution of all the variables. This means 'breaking down' the joint probability distribution into smaller marginal and conditional components. This decomposition reflects the (recursive) structure among the variables.[4] This is also called, following Blalock, the 'conceptual model'. Consider Gaumé and Wunsch (2010) again. Simplifying things quite a lot, the authors come up with a conceptual model where 'Self-rated health', the response variable (effect), directly depend on 'Education', 'Alcohol consumption', 'Locus of control', 'Psychological distress', and 'Physical health'. In their conceptual model, there are also indirect paths, for instance from 'Social support' to 'Self-rated health' *via* 'Psychological distress'—see fig. 1.
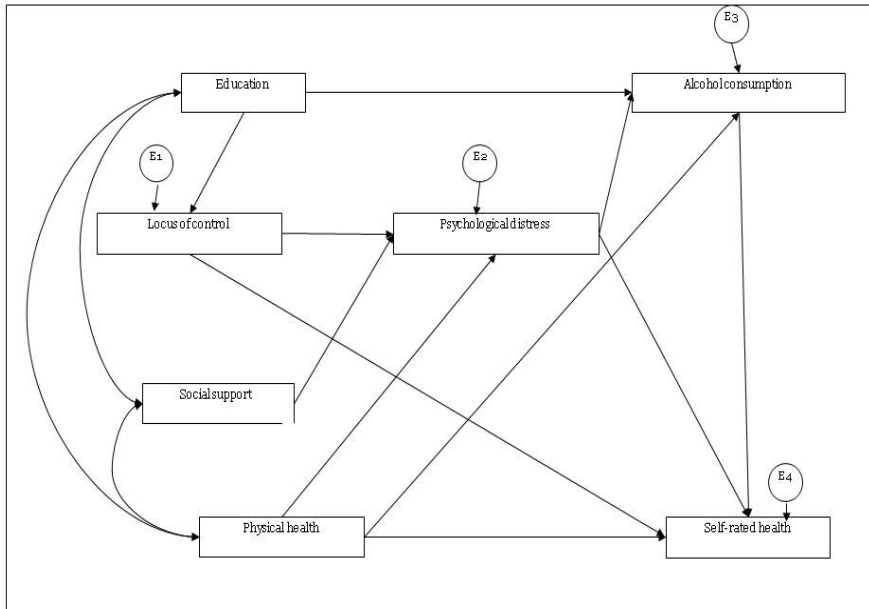

Figure 1: Determinants of self-rated health in the Baltic countries, 1994-1999.

The *third* step is to translate a conceptual model into an operational model. This means choosing the variables that can be directly measured or proxies for them, choosing the statistical model and the methods of testing. For instance, Gaumé and Wunsch (2010) fitted the model for four age groups (18–29, 30–44, 45–59, 60+), for both genders, for local individuals and for foreigners (mainly Russians). The authors evaluated the model through Bayesian structural equation modelling using a Monte Carlo Markov Chain (MCMC) procedure.

---

[4]It is customarily assumed that causal models are recursive, that is feedback loops are not permitted. I do not enter here the debate on the plausibility of such assumption, nor the debate on the methods to deal with non-recursive models.

Once the operational model is in place, the *fourth* step consists in testing the model for stability or invariance: *what* invariance is the matter of controversy with manipulationist theorists and I thoroughly address this issue later.[5] In Gaumé and Wunsch (2010) it turned out that the determinants taken into account (alcohol consumption, physical health, psychological health, psychological distress, education, locus of control, and social support) had a remarkable stable impact on self-rated health across the different Baltic countries, across the time-frames analysed, across gender, ethnicity, or age group.

Variational epistemology   In the previous section, I considered Horn (b) of the dilemma: if we grant a *charitable* reading to the manipulationist approach, it turns out that the rationale underpinning causal reasoning is misleading and disingenous. In a nutshell, I am about to argue that the rationale underpinning causal reasoning—both in experimental and observational methods—lies in the notion of *variation*, not manipulation (for a thorough discussion, see also Russo (2006, 2009)).

To understand what a rationale is and does, we need a brief recap on the philosophy of causality. In the philosophy of causality, three broad areas of investigation may be distinguished: metaphysics, epistemology, and methodology. The metaphysics of causality seeks to answer questions about what causality (or a cause) is. It is worth noting that conceptual analysis, in attempting to provide identity conditions for causal claims, or the ultimate content of various locutions such as 'A causes B', also contributes to answering questions akin to purely metaphysical ones. The epistemology and methodology of causality, instead, seek (i) to answer questions about how we know about causal relations and (ii) to develop or implement methods for discovery or confirmation of causal relations. It is worth noting that the border between epistemology and methodology is much more blurred than the border between metaphysics and epistemology-methodology.

The quest for a *rationale* of causality falls within the epistemology and methodology of causality and seeks to answer the following question. When we reason about cause-effect relations, what notion guides this reasoning? Is it regularity? Invariance? Production? Manipulation?

It is worth emphasising that rationale and identity conditions are not the same thing. A *rationale* is a principle, notion or concept underlying some decision, reasoning, modelling, or the like. *Identity conditions*, instead, are conditions under which a causal claim is true. According to manipulationist accounts, for instance, 'X causes Y is true' if, and only if, were we to manipulate X would yield changes to Y.

Let me now explain how the rationale of variation works. I give here just a taste of an argument from the causal modelling methodology presented above—the full argument, as well as other arguments, supporting the rationale of variation can be found in Russo (2009).

Causal modelling is regimented by a single rationale guiding model building and testing: the rationale of variation. For the sake of simplicity, consider the reduced form of a structural equation: $Y = \beta X + \epsilon$, where Y is the putative

---

[5]Invariance is not the only test performed in causal models. Causal models also need to pass tests about goodness of fit or about exogeneity. I am just keeping the discussion focused on the controversy with manipulationist theories.

effect, X the putative cause, β a parameter quantifying the effect of X on Y and ϵ represents the errors. The first question to be asked is whether the data set reveals meaningful co-*variations* between X and Y. If there are such meaningful co-variations, a second question arises: are those variations chancy or causal? In order to assess whether co-*variations* between X and Y are chancy or causal, we perform a number of tests, including (and perhaps most importantly) tests for invariance.

It is important to notice that the causal equation above can be given a *variational* and a *manipulationist* reading. However, whilst the former is more basic, the latter is derived. Let me explain further. At bottom, a structural equation is read as follows: variations in the putative cause X are accompanied by variations in the putative effect Y. How much Y varies in response to the variation in X is quantified by the parameter β. But this does not imply by all means that X has been manipulated. It could well be, as is typically the case in observational studies in social science, that statements about co-*variations* are based on calculated statistical correlations between the variables. The manipulationist reading is then derived from this basic variational reading as follows. In an experimental setting, manipulations on X make X varying, such that Y varies accordingly. In a controlled experiment, therefore, co-variations in X and Y are due to manipulations, unlike in observational studies.[6]

This is all to say that variation not only guides causal reasoning in observational settings, but does so also in experimental ones. Notably, the variations we are interested in are exactly those due to the manipulations on the putative cause. In this sense, variation is a *pre*condition to other notions, notably to manipulation. This does not imply that there is no role left to manipulation, though. Manipulation is still a *tool* to find out what causes what, *when* it can be actually performed, but not always.

Invariance   The last issue to address is *what* kind of invariance is needed in order to establish whether change-relating generalisations are causal or not. Invariance, I argue, does not require interventions *strictu sensu*. This means that manipulation is not a necessary tool to establish what causes what. Instead, what is required in absence of manipulation is that the relation between the putative cause and effect(s) remains sufficiently stable across different partitions of the data set or across similar populations analysed in the same study.

In Gaumé and Wunsch (2010), no manipulation is performed. Causal assessment is instead made through testing the stability of the putative causal relationship across different 'portions' of the data set. The different 'portions' have to be carefully chosen. In fact, if we test invariance across sub-populations randomly sampled, we should indeed expect to find, approximately, the same values but with a larger confidence interval; consequently, this test wouldn't be terribly useful.[7] Instead, we should appropriately choose sub-populations,

---

[6]To be sure, there is also a counterfactual reading, which is, just like the manipulationist reading, derived from the basic variational one. Under the counterfactual reading, the equation says that *were* we to change X, Y *would* accorning change. Notice that testing invariance under the counterfactual reading is far from being a trivial task. Some (e.g. Psillos (2004)) have even come to the conclusion that the manipulationist account is, in this respect, parasitic on the existence of laws of Nature, which would justify why it is the case that Ohm's law turns out to invariant under (counterfactual) intervention.

[7]Thanks to Guillaume Wunsch for drawing my attention to this point.

for instance considering different age strata, or different socio-demo-economic characteristics, or different geographical regions, or different time frames.

Invariance tests whether the relation between the cause-variable and the effect-variable(s) in fact shows some 'stability of occurrence'. In other words, whether an empirical generalisation captures a relation between X and Y that, in a given data set, is 'regular enough'. Notice, however, that the kind of regularity hereby required is at variance with the 'traditional' Humean regularity. In fact, invariance is not a condition of regular succession of effect-events following cause-events, but a condition of constancy of the characteristics of the relation itself, notably of the causal parameters (for a discussion, see Russo (2009, ch. 4)).

So it seems that the manipulationists' requirement that empirical generalisations be invariant *under intervention* is, in nonexperimental contexts, pretty strong. The reason, as we have seen, is that in those cases we cannot intervene, and yet some form of invariance is required nonetheless. Some manipulationist theorists, apparently, agree. According to Woodward (2003, ch. 5.8), for instance, there are 'possible-cause' generalisations that state, at bottom, that the presence of a type cause C raises the probability of an effect of type E. One example used by Woodward is 'Latent syphilis causes paresis'. These 'possible-cause' generalisations are exactly the kind of generalisations established by means of causal models, routinely used in the special sciences. Here, the invariance requirement is weakened (Woodward, 2003, ch. 6.15 and 7.8): 'weak invariance' is not to test whether the generalisation would remain stable were we to intervene, but whether the generalisation is stable across subpopulations.

To illustrate, Woodward (2003) discusses a pioneer paper on the relations between smoking and lung cancer. Woodward notices that this paper was written in 1959, when detailed knowledge about the biochemical mechanism through which smoking produces cancer was still lacking. Thus, this study largely relies on epidemiological evidence—that is observational data—and only to a lesser extent on experimental studies of laboratory animals. Woodward then points out that the authors do not aim to formulate 'exceptionless generalisations' (i.e. laws); instead they establish a causal link between smoking and lung cancer because the relation turns out to be *invariant*. What kind of invariance? Exactly the kind of invariance discussed above, namely stability of the relationship across subpopulations. Let us read the passage (Woodward, 2003, p. 312):

> For example, the authors note that some association appears between smoking and lung cancer in every well-designed study on sufficiently large and representative populations with which they are familiar. There is evidence of a higher frequency of lung cancer among smokers than among nonsmokers, when potentially confounding variables are controlled for, among both men and women, among people of different genetic backgrounds, across different diets, different environments, and different socioeconomic conditions [. . .]. The precise level and quantitative details of the association do vary, for example, the incidence of lung cancer among smokers is higher in lower socioeconomic groups, but the fact that there is some association or other is stable or robust across a wide variety of different groups and background circumstances.

13

The difference between the account of invariance hereby offered from the one of the manipulationists is that invariance is *not* counterfactually defined, *nor* does it necessarily involve manipulation. In making this move I am not claiming originality, as this kind of invariance is currently employed by practising scientists, and was indeed envisaged by Woodward. My point in these discussions is that *non*-counterfactual invariance, that is invariance *not* based on manipulation is methodologically more fundamental.

To sum up. In the light of the difficulties faced by manipulationism, empirical generalisations need to be reassessed. I carried over from the manipulationist account the idea that empirical generalisations are change-relating relations between variables. I then proposed that the question of whether empirical generalisations are causal or not is settled by testing a condition of invariance that does not take manipulation as methodologically fundamental. It is in fact *not* manipulation that provides the rationale underpinning both observational and experimental methods for causal inference, but rather the notion of variation.

5   Conclusion

Manipulationist approaches hold the view that information about the outcomes of interventions are needed for a variety of scientific purposes, e.g. causal assessment or explanation. This paper narrowed down the scope to the role of manipulation for causal assessment rather than explanation. The solution given by manipulationist theorists is that an empirical generalisation between X and Y is *causal* insofar as it is invariant under intervention. That is, an empirical generalisation is causal if, and only if, the relationship between X and Y is invariant under intervention, and it is not disrupted as a consequence of such intervention.

I argued, however, that manipulationim is trapped in a dilemma. The manipulationist project can in fact be read in two ways. On the one hand (Horn 1), if the project is given a conceptual reading, then it appears to be unilluminating from a methodological point of view. On the other hand (Horn 2), if the project is given a methodological reading, then a second dilemma opens up. If methodological manipulationism is strictly interpreted (Horn (a)), then it fails to provide the methodology for observational studies. Or, if it is charitably interpreted (Horn (b)), then the requirement of manipulation becomes so vague and weak as to be not only unilluminating, but also misleading in providing a rationale for causal reasoning in either experimental or observational studies.

In the light of the previous discussion, I reassessed empirical generalisations. Empirical generalisations, I argued, are indeed change-relating (or, *variation*-relating) relations—this is the core of agreement with manipulationist theorists. The importance of such characterisation of empirical generalisation is that it mirrors the rationale underpinning both experimental and observational studies: it is the notion of *variation*—not manipulation—that guides causal reasoning. I also agree that for empirical generalisations to be causal, they have to be invariant. Yet, the kind of invariance needed does not take manipulation as methodologically fundamental.

The broad conclusion to be drawn is that manipulation is *not* the building block of causal assessment. Manipulation is certainly a good *tool, when* it can be

performed, but not always. In other words, there is still room for sound causal assessment in the absence of manipulation. Granted, it is no surprise that, ceteris paribus, manipulations give us higher confidence in causal assessment. But the ceteris paribus clause is important. Well designed observational studies may deliver more reliable results than poorly designed controlled experiments.

References

Baumgartner, M. (2009). Interventionist causal exclusion and non-reductive physicalism. *International Studies in the Philosophy of Science*, 23(2):161–178.

Birkland, T. (2010). *An introduction to the policy process: theories, concepts, and models of public policy making.* M.E. Sharpe, third edition.

Blalock, H. M. (1961). *Causal inferences in nonexperimental research.* The University of North Carolina Press.

Cartwright, N. (2007). *Hunting causes and using them: approaches in philosophy and economics.* Cambridge University Press, Cambridge.

Eells, E. (1991). *Probabilistic causality.* Cambridge University Press, Cambridge.

Fisher, R. A. (1935). *The design of experiments.* Oliver & Boyd, Edinburgh, 1st edition.

Gaumé, C. and Wunsch, G. (2010). Self-rated health in the baltic countries, 1994–1999. *European Journal of Population.*

Glennan, S. (2010). Mechanisms, causes, and the layered model of the world. *Philosophy and Phenomenological Research.*

Hausman, D. (1997). Causation, agency, and independence. *Philosophy of Science*, 64(4):S15–S25. Supplement.

Hausman, D. and Woodward, J. (2004). Manipulation and the causal markov condition. *Philosophy of Science*, 71(5):846–856.

Morgan, S. L. and Winship, C. (2007). *Counterfactuals and causal inference.* Cambridge University Press, New York.

Mouchart, M. and Russo, F. (2010). Causal explanation: recursive decompositions and mechanisms. In McKay-Illari, P., Russo, F., and Williamson, J., editors, *Causality in the sciences.* Oxford University Press.

Mouchart, M., Russo, F., and Wunsch, G. (2009). Structural modelling, exogeneity and causality. In H. Engelhardt, H-P Kohler, A. P., editor, *Causal Analysis in Population Studies: Concepts, Methods, Applications*, chapter 4, pages 59–82. Springer, Dordrecht.

Pearl, J. (2000). *Causality: models, reasoning, and inference.* Cambridge University Press, Cambridge.

Psillos, S. (2004). A glimpse of the secret connexion: harmonising mechanisms with counterfactuals. *Perspectives on Science*, 12(3):288–319.

Russo, F. (2006). The rationale of variation in methodological and evidential pluralism. *Philosophica*, 77, Special Issue on Causal Pluralism:97–124.

Russo, F. (2009). *Causality and causal modelling in the social sciences. Measuring variations.* Methodos Series. Springer, New York.

Russo, F. (2011a). Correlational data, causal hypotheses, and validity. *Journal for General Philosophy of Science*, In press.

Russo, F. (2011b). Explaining causal modelling. or, what a causal model ought to

explain. In Sinigaglia, C. and D'Agostino, M., editors, *Proocedings of the SILF Conference 2007*.

Russo, F., Wunsch, G., and Mouchart, M. (2008). Potential outcomes, counterfactuals, and structural modelling: Causal approaches in the social sciences. In *European Population Conference*, Barcelona.

Skyrms, B. (1980). *Causal necessity: a pragmatic investigation of the necessity of laws.* Yale University Press, New Haven.

Strevens, M. (2007). Essay review of woodward, making things happen. *Philosophy and Phenomenological Research*, 74:233249.

Strevens, M. (2008). Comments on woodward, making things happen. *Philosophy and Phenomenological Research*, 77:171192.

Thompson, R. P. (2011). Causality, theories, and medicine. In McKay-Illari, P., Russo, F., and Williamson, J., editors, *Causality in the sciences*. Oxford University Press.

Waters, C. K. (2007). Causes that make a difference. *The Journal of Philosophy*, CIV:551 – 579.

Woodward, J. (2002). What is a mechanism? a counterfactual account. *Philosophy of Science*, 69:S366–S377.

Woodward, J. (2003). *Making things happen: a theory of causal explanation*. Oxford University Press, Oxford.

Woodward, J. (2010). Causation in biology: stability, specificity and the choice of levels of explanation. *Biology and Philosophy*, 25:287–318.

Woodward, J. and Hitchcock, C. (2003). Explanatory generalizations, part i: A counterfactual account. *Nos*, 37(1):1–24.

Wunsch, G. (1988). *Causal theory and causal modelling*. Leuven University Press, Leuven.

Wunsch, G. (2007). Confounding and control. *Demographic Research*, 16:15–35.

Wunsch, G., Russo, F., and Mouchart, M. (2010). Do we necessarily need longitudinal data to infer causal relations? *Bulletin de Méthodologie Sociologique*, 106(1):1–14.